Methodology in Phonological Acquisition

Assessing the Joint Development of Speech Perception and Production

Tania S. Zamuner and Elizabeth K. Johnson

University of Ottawa / University of Toronto, Mississauga

1. Introduction

The emergence of speech perception and language production is integrally related, with the development of the perceptual system in many ways setting the stage for the development of the production system. Surprisingly, despite the necessary link between these two systems, they have traditionally been studied independently of one another. The lack of work focused on understanding how language perception and language production develop in concert is in part due to disciplinary boundaries. Classic work in the area of developmental perception has been most often carried out by psychologists, whereas classic work in the area of developmental production has been most often carried out by linguists. Communicating findings from one field to the other has only recently become commonplace as a growing number of researchers have been trained in both fields. Methodological challenges present another impediment to the integrated study of language perception and production. Many classic methodologies for studying developmental speech perception are best suited for use with infants, and rely on cross sectional designs that reveal typical developmental patterns in groups of children. Many classic production methodologies, on the other hand, are best suited for longitudinal studies of individual or small samples of young children. The main goal of this chapter is to provide an overview of some recent research methodologies that are likely to play an important role in furthering this area of research. Note that this chapter is not an exhaustive review of all language testing methodologies.

Infant speech perception research has focused on development in children between birth and about 20 months of age. The bulk of language production research, on the other hand, has focused on language learners at the onset of word production, i.e. 18-20 months and older. (Though see work on babbling and the transition to speech, e.g., Vihman, DePaolis and Keren-Portnoy, 2009, and other work tracking phonological development in children starting at 1 year of age, such as Demuth, Culbertson and Alter, 2006, Fikkert, 1994 and Levelt, 1994). Infant speech perception research has demonstrated that infants know a great deal about the phonological patterns of their language long before they produce their first words (Curtin and Hufnagle, 2009; Mani, this volume). A dramatic demonstration of infants' early sensitivity to phonological patterns comes from the word segmentation literature. During the latter half of the first year of life, infants already begin using language-specific knowledge of how words typically sound to locate likely word boundaries (e.g. Johnson, 2008; Johnson and Jusczyk, 2001; Jusczyk, Houston, and Newsome, 1999; Mattys and Jusczyk, 2001). At this young age, infants appear to have highly specified perceptual representations of words. Even single segment mispronunciations in unstressed syllables are readily detected by 7.5 to 10-month-olds (Johnson, 2005; Jusczyk and Aslin, 1995). Some have come so far as to argue that lexical representations are likely over-specified early in development (e.g. Houston and Jusczyk, 2000; Singh, Morgan, and White, 2004, though see van Heugten and Johnson, 2009).

Despite their apparent perceptual sensitivity to phonetic detail, children make systematic production errors as they produce words (Fikkert, 2007). For example, children learning Dutch make errors in producing word-initial voiced segments, such as mispronouncing dier 'animal' as tier (van der Feest, 2007) The majority of developmental speech production studies do not simultaneously examine perception, perhaps because of the assumption that child perception is adult-like. Classic studies showing children do not accept their own mispronunciations as acceptable utterances certainly support this view. For example, although a child says sip for ship, they will not accept sip for ship when produced by an adult (Smith, 1973). This indicates that the child can perceive the phonological differences between their (inaccurate) output and the adult target form. Another feature of child language that suggests children's perception is adult-like comes from widespread changes in their articulatory development. When a child's production abilities mature and they are able to produce segments previously produced incorrectly, children correctly produce the target segments in all contexts, without needing to rehear all instances. For example, take a child who produces the /pl/ in please incorrectly as pease. As the child develops, they are able to correctly produce the /pl/ in please along with other words that contain /pl/, such as play. This is only possible if the child had accurately perceived the /pl/ sequence and had correctly stored representations of these targets (though see literature on lexical diffusion[1]). The absence of analyses on development speech perception in early childhood has also stemmed from methodological considerations. For example, many studies of phonological development have based their analyses and conclusions on longitudinal corpus data (Ohala, 2008). While these analyses provide insights into the development of a phonological system across a single child or group of children, because they are production-based, they do not allow for analyses into children perception.

For these and other reasons, there is a gap in our understanding of phonological development during learners' transition to speech. Although many production studies assume that children's perception is accurate and/or that children's production errors are not the result of perceptual errors (cf. Ohala, 1999); research shows that children's perception is in fact, not entirely adult-like (Walley, 2005)[2]. Examining both perception and production at the emergence of language production allows one to more precisely characterize phonological acquisition. For example, a study of speech production would be enhanced with a simultaneous examination of the time course of speech perception. The perception study could provide additional and crucial information about the phonetic detail children's early phonological and lexical representations, how language is being processed, learned and represented. A study by Sundara, Demuth and Kuhl (2008) addressed these types of issues by simultaneously evaluating children's perception and production of 3rd person singular –s. Results revealed differences in performance on the two tasks: children who had a familiarity preference for grammatical sentences in the perception task had lower production accuracy scores, while children who preferred to listen to novel, ungrammatical sentences had higher production accuracy scores.

The goal of this chapter is to give an overview of a few select behavioural methodologies that are suitable to experimentally assess developmental speech perception (Anticipatory Eye Movement Paradigm) and production (Elicitation tasks and Non-Word Repetition Tasks) in children who are at the early stages of lexical development; i.e., children who are approximately 20 months and older. While there is a large amount of variability when children produce their first words and many children will produce their first words before this stage, we focus on the age at which one can reasonably gather production data from young children who are being assessed in a single visit to an experimental laboratory. Children aged 20 months have an average expressive vocabulary size of around 50 words (Fenson et al., 1997, cited in Dale et al., 1998). Even though it is possible to do production studies with children younger than 20 months (Ohala, 1999), in our experience, this age is at the lower end from which one can successful complete experiments using elicitation and non-word repetition.

2. Methods to Assess Developmental Speech Perception

As noted above, research on developmental speech perception examines acquisition from early infancy, starting at birth. Some of the methods commonly used to research speech perception in young infants, such as the Headturn Preference Procedure (Fernald, 1985, Kemler Nelson, Jusczyk, Mandel, Myers, Turk, and Gerken, 1995), are optimally used with younger infants. While the HPP can be used with 20-month-olds, older children can become bored with this procedure, leading to high attrition rates. Moreover, this paradigm does not typically provide a very sensitive measure of individual variation, making it less than ideal for examining the relationship between perception and production. One of the most recent techniques developed to examine brain activity in neonates and young infants is near-infrared spectroscopy (NIRS) (Aslin and Mehler, 2005; Mehler, Gervain, Endress and Shukla, 2008; Meek, 2002). NIRS has been applied to examined language development (Gervain, Macagno, Cogoi, Peña and Mehler, 2008; Peña et al., 2003), sensory processing (Bortfeld, Wruck, and Boas, 2007), prosody (Homae, Watanabe, Asakawa, and Taga, 2006; Homae, Watanabe, Nakano, and Taga, 2007), and language processing in bilinguals (Kovelman et al., 2008). This paradigm may be promising way to test developmental speech perception in toddlers, yet it is still in development and it is yet unknown how sensitive the paradigm will be at targeting specific issues in phonological acquisition.

Testing developmental speech perception with toddlers is somewhat tricky depending on the methodology that one uses. It can be difficult to obtain explicit judgments from young children because they often do not understand the instructions and find it difficult to focus. Consider a picture naming task used to assess discrimination of minimal pairs, such as a place of articulation contrasts between /d/ *doll* and /b/ *ball*. When asked, "*Point to the doll*", 20-month-olds will often simultaneously point to both pictures on the page, making it difficult to assess their perceptual abilities. Even when using a picture pointing methodology where pictures are presented on a computer, young children may enjoy the study, but do not perform well (Parisse and Soubeyrand, 2003). In short, picture pointing tasks and tasks that require children to provide an explicit answer may give too conservative an estimate of children's phonological knowledge (e.g., Barton, 1980; Brown and Matthews, 1997; Shvachkin, 1948/1973) as compared to on-line (implicit) measures (Fennell and Werker, 2003).

An established technique that can be used to assess speech perception in word forms with young children is the Intermodal Preferential Looking Procedure (IPLP) (Hirsh-Pasek and Golinkoff, 1996; Hollich, Hirsh-Pasek, and Golinkoff, 2000; Lew-Williams and Fernald, 2007; Houston-Price, Mather, and Sakkalou, 2007). The IPLP has received close attention in recent methodology overviews (Fernald, Zangl, Portillo, and Marchman, 2008; Gerken, 2009; Johnson and Zamuner, 2010), therefore, we will only mention this methodology in brief. In the IPLP, children are presented with trials containing two pictures shown on a visual display. For example, pictures of minimal pairs like *ball* and *doll*. Participants are asked to look at a particular picture, '*Look at the ball*'. Experimental analyses examine participants' proportion of looking time to the target picture (*ball*) versus the distractor (*doll*), and the relative speed to which participants look at the target and distractor. Recall that a limitation of picture pointing with very young children is that they will often point to both pictures when asked to "*Point to the ball*." In the IPLP task, children may *look* at both pictures, however, the methodology also allows one to measure the proportion of looking time to either picture and the speed of looking, providing a more sensitive measure of young children's phonological knowledge.

In the field of phonological development, findings based on this methodology have had a large impact on our understanding of children's early phonological and lexical representations. In particular, results have illustrated that children's early lexical representations encoded much more phonetic detail than previous production studies have suggested (Swingley and Aslin, 2002). At 14-months, infants look longer to an object when it is produced correctly (*baby*) than misproduced (*vaby*). Though note that recent work using the same methodology has argued that early lexical and phonological representations are underspecified (van der Feest, 2007). While the IPLP is a powerful technique for investigating phonological acquisition at the onset of language production, one shortcoming is that it requires pictureable objects, from which looking times are measured. If the appropriate real-word stimuli are not familiar to young children, it may be useful to use non-word stimuli (though this methodology can also provide insights into word learning mechanisms, Byers-Heinlein and Werker, 2009). In the following section, a new paradigm called the Anticipatory Eye Movement Paradigm is described, a paradigm that does not require picturable (known) objects to test speech perception.

2.2 Anticipatory Eye Movement Paradigm (AEM)

A new and promising methodology that may help researchers better understand the perception-production link is the Anticipatory Eye Movement Paradigm (AEM) (Aslin and Fiser, 2005; McMurray and Aslin, 2004). This methodology was developed to study categorization in language learning, although its foundations are in the visual expectation paradigm (VEP) (Haith, Hazan and Goodman, 1988). This methodology and variations of it have been used in different areas of cognitive development, such as spatial representations (Kaufman, Gilmore and Johnson, 2005), object concepts (Johnson, Amso and Slemmer, 2003), cognitive control abilities (Kovács and Mehler, 2009a). It has also been used in studies of language, such as on-line sentence processing (Altmann and

Kamide, 2007; Nation, Marshall and Altmann, 2003). A recent study by Kovács and Mehler (2009b) used this technique to investigate monolingual and bilingual 12-month-old infants' sensitivities to structural regularities in speech. They found that bilingual infants were better at learning multiple speech structures as compared to the monolingual infants (trisyllabic speech items with an AAB or ABA structure). The suggestion is that bilingual infants are more flexible learners and this enables them to learn multiple languages.

In AEM, infants are trained to anticipate different categories of acoustic (and/or visual) stimuli on either the left or right side of a visual display. Importantly, training stimuli unambiguously belong to 1 of 2 categories and there is typically 100% predictable relationship between what a child hears (and/or sees) and where the visual reinforcer will appear. For example, stimulus A is associated with the left side of the visual display, and stimulus B is associated with the right side of the visual display. See Figure 1 which displays the time course of a trial taken from McMurray and Aslin (2004). After infants have learned that stimuli type A leads to visual reinforcement on the left side and stimuli type B leads to visual reinforcement on the right side, they begin making anticipatory movements to the left or right side of the screen as soon as the target training stimuli are presented. During test, novel stimuli are played and infants' eye movements are recorded. Test stimuli differ from the training stimuli in that they do not necessarily unambiguously belong to one category or another. For example, in one study, infants were trained on two endpoints of the *ba*/*pa* continuum, and then tested on new exemplars with intermediate voice onset times (VOTs) (McMurray, Spivey, and Aslin, 2000). AEM has also been used to show that infants are able to identify words across changes in pitch, but not duration

(McMurray and Aslin, 2004). When used in discrimination paradigms, the AEM has been shown to be more sensitive than classical behavioural tasks. For example, Albareda-Castellot, Pons, Sebastián-Gallés (2008) adapted the AEM paradigm so that all trials had visual reinforcement. Infants' anticipatory eye movements across the course of the experiment were compared to assess infants learning. Albareda et al. found that 8-month-old Spanish-Catalan bilinguals were able to discriminate Catalan vowel contrasts, whereas studies using familiarization methodologies with the same aged and language background infants did not find discrimination (Bosch and Sebastián-Gallés, 2003).

AEM studies can be designed in a number of ways. As stated before, studies typically have a training phase and a test or generalization phase. Experiment designs may require a predetermined number of training trials before the test phase (McMurray and Aslin, 2004). Alternatively, the length of the training phases may be tailored to individual participants. This is due to the recent development of software that analyzes infants' anticipatory looking over the course of training. For example, programs can calculate an on-line measure of the number of correct trials with correct anticipatory looking. Based on these measures, it is possible to stipulate that once participants have reached a preset criterion of looking, they proceed to the test phase where generalization trials are presented with novel and potentially ambiguous stimuli. As with the IPLP, the dependent measure can vary from the proportion of looking time to the either side of the screen, the relative speed to which participants look to either side of the screen or with time course analyses. Analyses of the training phase and test phase will depend on the experiment design. Importantly, analyses of the training phase should examine whether participants make the correct anticipatory eye movements. Analyses of test trials will show how individual participants or groups of participants categorize or generalize new stimuli in comparison to the training stimuli.

The AEM is adaptable and is not limited to any specific age or linguistic groups; however, it is not appropriate for infants under 2 months, as controlled eye movement is still developing around this time. In our work, we have begun using this paradigm with 12 to 20 month olds, and are finding the task to be well suited to this age range. AEM studies are best presented using eye-tracking systems because they allow for precise measurements and timing of infants' eye movements. Moreover, with the use of eye-tracking techniques, the procedure easily allows for reaction time measures. A major disadvantage in using AEM is the relative newness of the paradigm. Standard experimental designs and analyses have not yet been established. Although few studies have been published using this paradigm with phonological development, AEM studies have a long history in other areas of cognitive development.

3 Methods to assess development speech production

Experimental research in phonological speech production has received less attention than in developmental speech perception. Typically, production studies have focused on evaluating the development of individual or groups of children's phonological systems. This approach stems in part from Smith's landmark study (Smith, 1973), in which he provided a comprehensive analysis of his son's phonological development. Large-scale corpora studies are labour intensive, though database sharing through CHILDES (MacWhinney, 2000) and the development of corpora tools such as PHON, an open-source program to manage phonological corpora (Rose et al., 2006) are making it attainable for researchers to investigate phonological development using large scale, longitudinal corpora. For a review on issues pertaining to the development of corpora suitable for examining language development see Demuth (2008).

There are a number of techniques that can be used to assess phonological development in production studies with children. For example, word-games require participants to manipulate parts of words allowing researchers to examine participants' knowledge of words' segmental and syllable structure (Fallows, 1981; Treiman and Danis, 1988; Zamuner and Ohala, 1999). To illustrate this, take a pause-insertion task. Here participants are asked to repeat words with a pause between syllables, such as *ti**g**er* is repeated as *ti**g***(pause)*er*, *ti*(pause)***g**er*) or *ti**g***(pause)***g**er*. Analyses of children's responses determine whether participants associate the medial 'g' with the first syllable, second syllable, or as ambisyllabic. The association (syllabification) of the medial consonant is influence by factors such as vowel quality and stress. Another task used with young children is the Wug Task, which was designed to assess learners' productive morpho(phonological) knowledge (Berko, 1958). Participants are given novel words (*This is a wug*.) and asked to produce a morphological variant (*Now there is another one*. *There are two of them*. *There are two* ____. Answer: *wugs*.). While methodologies using word-games and other types of manipulations are informative, they are also difficult for toddlers to complete. Therefore we focus on two methodologies that have been successfully used with children under the age of two: elicitation and non-word repetition. We also point the reader to Blom and

Unsworth (2010), for a discussion of issues in running experimental language studies with children.

3.1 Elicitation studies

When using corpus data, a common drawback is that in a typical recording session of spontaneous, naturalistic data, children will not produce enough words (or any words) containing the phonological structure under investigation. Take an example of naturalistic data on young children's early morpho-phonological voicing alternations in Dutch from the CLPF database (CHILDES, Fikkert, 1994; Levelt, 1994). In Dutch, voicing alternates between [t~d] in pairs like [brot~brodən] 'bread~breads'. Although the database contains over 20,000 utterances collected from 12 children acquiring Dutch, there are only 9 types that have the targeted voicing alternations (Kerkhoff, 2007: 132). To address this limitation in using naturalistic data, one of the simplest tasks to use with young children is elicitation (Ohala, 2008). In this methodology, children are prompted for words that are targeted for their phonological characteristics. Elicitation tasks are based on the assumption that the patterns of correct and incorrect productions reflect learners' phonological knowledge, language acquisition mechanisms, and language representations. Using elicitation task, researchers have examined a wide range of phonological patterns. For example, studies have investigated learner's knowledge of segmental structure, using words with different clusters to examine children's error substitutions (Kirk, 2008), and studies have examined the acquisition of prosodic structure, using multisyllablic words to study truncation patterns (Kehoe and Stoel-Gammon, 1997).

There are a variety of tasks that can be designed to elicit productions of specific words. A very successful method is the picture-naming task. In our experience, this is more successful when pictures are presented on the computer as children find it very engaging. It is also possible to intersperse entertaining video trials, to maintain their attention, which will vary depending on the age of a child. Another major advantage of using a computer is that it keeps children in a specific location within the room, allowing for more reliable audio and video recordings. This can be essential for later acoustic analyses and/or transcriptions. Children will regularly place their hands in their mouth, and if these are captured on video, one is able to exclude these productions. See the CHILDES website for a complete discussion on how to make audio and video recordings (childes.psy.cmu.edu).

Many issues can influence children's production, and it is essential to consider what factors to control in the set of experimental stimuli and methodology. For example, recent work by Edwards and Beckman (2008) compared children's production of word-initial segments in Cantonese, English, Greek and Japanese. Children's segmental accuracy was effected by the frequency of the neighbouring phoneme context (initial segments were more accurate when followed by a frequent vowel than an infrequent vowel), word length (initial segments were more accurate in shorter words than longer words), and prosodic factors such as word stress and pitch accent (initial segments in Japanese are produced more accurately when they are in syllables with high tone than low tone). Other factors such as utterance position can influence the accuracy of children's production. For example, children's production of 3rd person singular morphemes (-s in *looks*) is produced more often in utterance-final than utterance-medial position (Song, Sundara, and Demuth, 2009). Another factor is word frequency; children are better at producing high frequency words and frequently occurring phonological patterns (Zamuner, 2003). Therefore, before deciding on the set of experimental items, it is often useful to calculate the distribution of phonological patterns in the ambient language in an appropriate corpus (e.g., a corpus of child directed speech). Controlled experimental studies can help factor out some of these influences on children's production, which are difficult to control for in analyses of corpora data.

Elicitation tasks can be very successful with young children. The only limitation is that participating children must be comfortable enough in the experimental setting to produce speech. Young children are often shy in new situations; therefore, it is useful to first use a task or play a short game that does not require verbal responses. In some cases participants will not produce any spontaneous responses, thus, it is useful to prepare an imitation task for the same stimuli. This may be part of a between-subjects condition where imitation can be compared to spontaneous productions. Imitation can be informative because it can help interpret elicited productions. For example, using a picture-naming task, Zamuner, Kerkhoff and Fikkert (2008) found that 3 ½ year old Dutch-learning children do not reliably produce a medial voicing contrast in certain morphological conditions. In bi-morphemic contexts, children produce both medial 't' in *petten* 'caps' and medial 'd' in *bedden* 'beds' as 't', *petten* and *betten*, respectively. In mono-morphemic conditions, children accurately produce a voicing contrasts, the medial 't' in *water* and the medial 'd' in *poeder*' are produced correctly. However, when the same children were tested on an imitation task, a different result was found; children were equally accurate at imitating medial ‘t’ and medial ‘d’ in both morphological contexts. This provides evidence that children at this age do not have difficulty producing a medial voicing contrast, but that other lexical factors influence their performance in the picture naming task.

When designing an elicitation task, there will be a limitation in the number of words that you can reasonably expect children to produce in a single experimental session, which will change depending on the specific age or age range of the participants. To cope with these limitations, one could use a between subjects design or have multiple testing sessions with the same child. When the study includes very young children, extensive piloting is helpful to determine whether children are likely to be familiar with the items and pictures. To help solicit items, specific frames should be prepared and used with all children for consistency. It may also be necessary to include filler items. Research has shown that over sampling of a specific word shape or phonological pattern may prime children to produce words in a specific way, lead to a higher number of errors than found in spontaneous speech. For example, an analysis of children’s overgeneralization errors found that children produced more overgeneralization errors in an experimental setting than found in spontaneous speech. In the experimental study, children were provided with a verb stem and asked to produce it in the past tense, as in ‘I will drink my milk. I already _____ my milk.’ (Kuczaij, 1978; Marcus, Pinker, Ullman, Hollander, Rosen and Xu, 1992). In this context, children were more likely to overgeneralize ‘drinked’ than in spontaneous speech samples. The use of fillers can help offset these potential priming effects.

Once the data are collected, they will need to be transcribed and coded. Protocols for phonetically transcribing child production data vary across researchers and experimental labs. For example, variation is found in the use of trained phonetic transcribers who may or may not be native speakers of the language, how naïve the transcribers are to the experiment purpose, and the amount of transcribed data that is checked for inter-transcriber reliability. The level of phonetic transcription also varies (narrow versus broad) depending on the goals of the study, though it should be noted that there is less reliability across transcribers on narrow phonetic transcriptions. Many of these issues have important implications for transcribed data. See Edwards and Beckman (2008) for a recent discussion on how to deal with issues of phonetic transcriptions, such as influences of transcribers' native-language bias in their transcription of non-native language phonemes. Some of these issues can be addressed by including acoustic analyses of children's production. These types of analyses may reveal covert contrasts that exist in children's speech, but that are not audible to adult transcribers (Buder, 1996; Scobbie, Gibbon, Hardcastle, and Fletcher, 2000). Alternatively, acoustic analyses may be the primary way in which the data are coded, depending on the study's goals.

Data analyses will depend on the goals of the study. An experimental goal may attempt to determine the sound patterns that children are able to produce at different stages. For example, if one is investigating whether there are prosodic interactions in children's segmental productions, one may compare children's accurate production of word-initial versus word-final segment, or the different types of segments produced accurately in these positions (such as targets with labial, coronal or dorsal place of articulation) (e.g., Beers, 1995). It may also be informative to do an error analysis, to see whether children's errors show specific patterns of results. For example, when young children first attempt consonant clusters, they typically will delete one segment. Analyses may examine whether children reduce the first or second member of the cluster, e.g., does *pretty* reduce to *pitty* or *ritty*. These types of analyses have revealed that children's cluster reductions adhere to the well-formedness of syllable sonority (Jongstra, 2003). Alternatively, analyses may examine participants' segmental substitutions. For example, these types of analyses have revealed that children's segmental substitutions result in clusters that share place or manner of articulation, e.g., *ducks* becomes *duts* (Kirk, 2008).[3] Other types of error analyses may examine whether children's substitution errors are towards the more frequent segmental substitutions or more phonological features. Also it may be possible to collect reaction time measures from children's productions using a voice key, which provide an automatic and electronic measure (rather than an off-line, manual evaluation) of the time between stimulus presentation and the onset of speech production (Tyler, Tyler, and Burnham, 2005). In this case, one might compare reactions for children's correct productions versus incorrect productions of initial segments.

An unavoidable problem that is frequently encountered when working with young children is empty data cells. That is, young children may not produce a specific item for various reasons. This can usually be addressed by taking the proportion of correct responses for a condition. For example, consider an experiment designed to examine children's production abilities of place of articulation. In this hypothetical study, there are four words in the stimulus set to evaluate how children produce word-initial /b/: *bear*, *book*, *boat* and *bed*. Imagine a participant who correctly produces /b/ in *bear* and *book*, misproduces the /b/ in *boat*, and does not give a response for *bed*. For this participant, their accuracy score on word-initial /b/ would be .66 (2 out of 3 words with word-initial /b/ were produced correctly).

Although there are many advantages to using real-word stimuli, there are also potential limitations. With real word stimuli, it is often difficult to find enough appropriate items that are picturable or familiar to young children. The targeted sound pattern may occur most often in verb or adjectives, making it difficult to elicit spontaneous productions of these words as compared to nouns. For example, there are very few common English words known to children that in end /v/ as in *love*, whereas it is very easy to find nouns that end in /t/, as in *cat*. The frequency of individual lexical items also has an effect on how accurately children produce words. For example, Zamuner (2003) found that young children are quite accurate at produced final 'th' in the word *bath*, even though 'th' is a late acquired segment and is difficult for young children to produce. Other processes such as lexical diffusion may influence children's performance. Lexical diffusion refers to how sound changes evolve in children's developing phonological stems. Children's acquisition of fricatives has shown different patterns of development depending on word frequency and where the target fricative occurred within a word's position (Gierut and Storkel, 2002). Given these types of potential limitations of using real word stimuli, it can be hard to design controlled experiments to test phonological development of specific phenomena.

3.2 Non-word repetition tasks

One methodology that circumvents these problems is the non-word repetition task (NWRT). The NWRT has been widely used to assess the development of children's phonological and lexical representations, speech perception, articulation and memory (see Coady and Evans, 2008 for a recent review). For example, parallel to the studies that have used elicitation to study the acquisition of consonant clusters, Ohala (1999) found in a non-word repetition task, that children's reductions of initial and final clusters were predicted by sonority. Initial clusters were reduced to produce a rise in sonority ([stig] → [tig]), whereas final cluster reduction more often to led to a minimal sonority descent ([dust] → [dus]). Similarly, studies looking at prosodic acquisition have examined truncation patterns in young children's production of multisyllablic words (Gerken, 1994).

In non-word repetition tasks, children are simply asked to repeat non-words that are controlled for various phonological properties (or other types of properties). Typically the methodology is used with children over the age of 3, although studies have been successful with participants under 2-years-of-age (Zamuner, 2003). Non-word repetition tasks are ideal for children at the beginnings of language production because they capitalize on 'echoism' (Jespersen, 1922/1964) or 'echolalia' (Guillaume, 1926/1971). That is, children imitate speech. In a research setting, young children who are presented with non-words will spontaneously repeat them, with needing explicit instruction to do so. As with the picture-naming task, we have had the most success presenting pre-recorded non-words over a computer. This also controls the acoustic properties of the non-word stimuli, so that all subjects are presented with the same tokens.

Like studies using elicitation, non-word repetition tasks require careful attention to stimuli design. The frequency of the sound components of non-words is an important factor to consider, as children are more accurate at producing frequent segments and segmental combinations (Coady and Evans, 2008; Munson, Kurtz and Windsor, 2005). Moreover, studies have revealed that young children are more accurate at producing the same sound depending on the frequency of the non-word components. Zamuner (2009) found that children are better at producing word-initial /p/ in non-words composed of high frequency patterns than low frequency patterns. A standard way in the field is to control non-word stimuli for their phonotactic probabilities, that is, the likelihood that a sound has to occur in a given word environment (Storkel, 2004). Stimuli are also typically controlled for their neighbourhood densities, which is a measure of the number of similar sounding words in the lexicon.

Many of the same considerations described in the previous section on elicitation apply to data transcription, data coding and data analyses for NWRT. A unique consideration in studies using non-word production is the treatment of real word responses. Young children may often produce a real word in response to a non-word. For example, in Zamuner, Gerken and Hammond (2004), a typical error in children's production of the non-word *bome*, was to substitute the final 'm' as 'n', producing *bone*. It is possible that children misperceived *bome* as *bone*, though it is also possible that children perceived it correctly and made a segmental production error. To address this problem, we have typically excluded real word answers when they are words known to young children – i.e., if the real word occurs in a corpus of speech known to young children. In this case, children's productions such as *bone* would be excluded, whereas a production of *bode* would not be.

4 Conclusion

In this chapter we have described methodologies suitable for investigating the simultaneous development of speech perception and production. The perceptual task we have focused on is the Anticipatory Eye Movement Paradigm. The two production tasks we have focused on are word elicitation (picture naming task) and the non-word repetition task. There are two broad approaches to combining these sorts of methodologies to examine the development of perception and production in tandem. To illustrate, take the case of voice onset time (VOT) acquisition. One approach would be to examine VOT perception and production in the same child. For example, following McMurray et al. (2000), one could devise a perception study using the AEM to examine a child's language specific voicing contrasts. In other words, to design a study that would establish the VOT boundary for an individual child. This same child could then be tested on a production study using real-words or non-word repetition task. Comparisons of the same child's system could be made to determine whether individual children's perceptual and production systems align. Another approach would be a similar type of study, but to compare group data on perception and production. Examining developmental speech perception and production in concert is an upcoming challenge for researchers. Yet the future is bright for research exploring the relationship between speech perception and speech production. We hope that the methodologies described here will provide useful tools for researchers interested in these overlapping and intertwined areas of language development.

References

Albareda-Castellot, B., Pons, R., Sebastián-Gallés, N. (2009). The acquisition of phonetic categories in bilingual infants: New data from a new paradigm. Manuscript under review.

Altmann, G. T. M., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. Journal of Memory and Language, 57, 502–518

Aslin, R. N., & Fiser, J. (2005). Methodological challenges for understanding cognitive development in infants. Trends in Cognitive Science, 9, 92-98.

Aslin, R. N., & Mehler, J. (2005). Near-infrared spectroscopy for functional studies of brain activity in human infants: promise, prospects, and challenges. Journal of Biomedical Optics, 10, 11009.

Barton, D. (1980). Phonemic perception in children. In G. H. Yeni-Komshian, J. F. Kavanagh & C. A. Ferguson (Eds.), Child Phonology: Volume 2 – Perception. New York: Academic Press.

Beers, M. (1995). The phonology of normally developing and language-impaired children. Amsterdam: Institute for Functional Research Into Language and Language Use.

Berko, J. (1958). The child's learning of English morphology. Word, 14, 150-177.

Blom, E. & Unsworth, S. (Eds.) (2010). Experimental Methods in Language Acquisition Research. Amsterdam: Benjamins ((Language Learning and Language Teaching series), forthcoming.

Bond, Z. S. (1999). Slips of the Ear: Errors in the Perception of Casual Conversation. (San Diego, CA: Academic Press.

Bortfeld, H., Wruck, E., & Boas, D. (2007). Assessing infants' cortical response to speech using near-infrared spectroscopy. Neuroimage, 34, 407-415.

Bosch, L., & Sebastian-Galles, N. (2003). Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of life. Language and Speech, 46(2-3), 217-243.

Brown C., & Matthews, J. (1997) The role of feature geometry in the development of phonemic contrasts. In S. J. Hannahs and M. Young-Scholten (eds.), Generative Studies in the Acquisition of Phonology (pp. 67–112). Amsterdam: John Benjamins Publishing Company.

Buder, E. H. (1996). Experimental phonology with acoustic phonetic methods: Formant measures from child speech. In B. Bernhardt, J. Gilbert, & D. Ingram (Eds.), Proceedings of the UBC International Conference on Phonological Acquisition (pp. 254–265), Somerville, MA: Cascadilla Press.

Byers-Heinlein, K., & Werker, J.F. (2009). Monolingual, bilingual, trilingual: Infants' language experience influences the development of a word learning heuristic. Developmental Science, 12, 815-823.

Coady, J. A., & Evans, J. L. (2008). Uses and interpretations of non-word repetition tasks in children with and without specific language impairment (SLI). International Journal of Language & Communication Disorders, 43, 1–40.

Curtin, S., & Hufnagle, D. G. (2009). Prelinguistic speech perception. In Bavin, E. (Ed.), The Cambridge University Press Handbook of Child Language (pp. 107-124). Cambridge: Cambridge University Press.

Cutler, A. (ed.) 1982. Slips of the tongue and language production, Mouton, Berlin.

Dale, P. S., Simonoff, E., Bishop, D. V. M., Eley, T. C., Oliver, B., & Price, T. S., et al. (1998). Genetic influence on language delay in two-year-old children. Nature Neuroscience, 1, 324– 328.

Demuth, K. (2008). Exploiting corpora for language acquisition research. In H. Behrens (Ed.), Corpora in Language Acquisition Research: Finding Structure in Data (pp. 199-205). Trends in Language Acquisition Research (TILAR) Series #6, John Benjamins.

Demuth, K., Culbertson, J. & Alter, J. (2006). Word-minimality, epenthesis, and coda licensing in the acquisition of English. Language & Speech, 49, 137-174.

Edwards, J., & Beckman, M. E. (2008). Methodological questions in studying consonant acquisition. Clinical Linguistics & Phonetics, 22, 937-956.

Fallows, D. (1981). Experimental evidence for English syllabification and syllable structure. Journal of Linguistics, 17, 309-331.

Feest, S. van der. (2007). Building a phonological lexicon: The acquisition of the Dutch voicing contrast in perception and production. Ph.D. Dissertation. Radboud University Nijmegen, The Netherlands.

Fennell, C. T., & Werker, J. F. (2003). Early word learners' ability to access phonetic detail in well-known words. Language and Speech, 46, 245-264.

Fenson, L., Pethick, S., Renda, C., Cox, J. L., Dale, P. S., & Reznick, J. S. (1997). Technical manual and user's guide for the MacArthur communicative development inventories: Short form versions. Unpublished manuscript, San Diego State University.

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. Infant Behaviour and Development, 8, 181-195.

Fernald, A., Zangl, R., Portillo, A. L., & Marchman, V. A. (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In Sekerina, I. A., Fernandez, E. M., & Clahsen, H. (Eds.) Developmental Psycholinguistics: On-line methods in children's language processing (pp. 97-135). Benjamins: Amsterdam.

Fikkert, P. (1994). On the acquisition of prosodic structure. Ph.D. Dissertation. The Hague: Holland Academic Graphics.

Fikkert, P. (2007). Acquiring phonology. In P. de Lacy (Ed.), Handbook of phonological theory (pp. 537-554). Cambridge, MA: Cambridge University Press.

Gerken, L. A. (1994). A metrical template account of children's weak syllable omissions from multisyllabic words. Journal of Child Language, 21, 565-584.

Gerken, L. A. (2009). Language Development. Plural Publishing.

Gervain, J., Macagno, F., Cogoi, S., Peña, M., & Mehler, J. (2008). The neonate brain detects speech structure. Proceedings of the National Academy of Sciences, USA, 105; 14222-14227.

Gierut, J. A. (2001). A model of lexical diffusion in phonological acquisition. Clinical Linguistics & Phonetics, 15, 19-22.

Gierut, J. A., & Storkel, H. L. (2002). Markedness and the grammar in lexical diffusion of fricatives. Clinical Linguistics & Phonetics, 16 115-134.

Guillaume, P. (1971). Imitation in children. Chicago: University of Chicago Press (Original work published in 1926).

Haith, M. M., Hazan, C., & Goodman, G. S. (1988). Expectation and anticipation of dynamic visual events by 3.5 month-old babies. Child Development, 59, 467-479.

Hirsh-Pasek, K., & Golinkoff, R. M. (1996). The intermodal preferential looking paradigm: A window onto emerging language comprehension. In D. McDaniel, C. McKee, & H. S. Cairns (Eds.). Methods for assessing children's syntax (pp. 105–124). Cambridge, MA: MIT Press.

Hollich, G., Hirsh-Pasek, K., & Golinkoff, R. (2000). Breaking the language barrier: An emergentist coalition model of word learning. Monographs of the Society for Research in Child Development, 65 (3).

Homae, F., Watanabe, H., Asakawa, T., & Taga, G. (2006). The right hemisphere of sleeping infants perceives sentential prosody. Neuroscience Research, 54, 276-280.

Homae, F., Watanabe, H., Nakano, T., & Taga, G. (2007). Prosodic processing in the developing brain. Neuroscience Research, 59, 29-39.

Houston, D. M., & Jusczyk, P. W. (2000). The role of talker-specific information in word segmentation by infants. Journal of Experimental Psychology: Human Perception and Performance, 26, 1570-1582.

Houston-Price, C., Mather, E., & Sakkalou, E. (2007). Discrepancy between parental reports of infants' receptive vocabulary and infants' behaviour in a preferential looking task. Journal of Child Language, 34, 701-724.

Jespersen, O. (1964). Language: It's nature, development, and origin. New York: Norton. (Original work published in 1922).

Johnson, E. K. (2005). English-learning infants' representations of word-forms with iambic stress. Infancy, 7, 95-105.

Johnson, E. K., & Zamuner, T. S. (2010). Using infant and toddler testing methods in language acquisition research. In E. Blom & Sharon Unsworth (Eds.), Experimental Methods in Language Acquisition Research. Amsterdam: John Benjamins.

Johnson, E. K. (2008). Infants use prosodically conditioned acoustic-phonetic cues to extract words from speech. Journal of the Acoustical Society of America, 123, EL144-EL148.

Johnson, E. K., & Jusczyk, P. W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. Journal of Memory & Language, 44, 1-20.

Johnson, S. P., Amso, D., Slemmer, J. A. (2003). Development of object concepts in infancy: Evidence for early learning in an eye-tracking paradigm. Proceedings of the National Academy of Sciences of the USA, 100, 10568-10573.

Jongstra, W. (2003). Variable and Stable Clusters: Variation in the Realisation of Consonant Clusters. The Canadian Journal of Linguistics, 48, 265-288

Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. Cognitive Psychology, 39,159-207.

Kaufman, J., Gilmore, R., & Johnson, M. H. (2006). Frames of reference for anticipatory action in 4-month old infants. Infant Behavior and Development, 29, 322-333.

Kehoe, M., & Stoel-Gammon, C. (1997). Truncation Patterns in English-Speaking Children's Word Productions. Journal of Speech, Language, and Hearing Research, 40, 526-541

Kemler Nelson, D., Jusczyk, P. W., Mandel, D. R., Myers, J., Turk, A. E., & Gerken, L. A. (1995). The headturn preference procedure for testing auditory perception. Infant Behavior and Development,18,111–116.

Kerkhoff, A. (2007). Acquisition of Morpho-Phonology: The Dutch Voicing Alternation. Ph.D. Dissertation. Utrecht University, The Netherlands.

Kirk, C. (2008). Substitution Errors in the Production of Word-Initial and Word-Final Consonant Clusters. Journal of Speech, Language, and Hearing Research, 51, 35-48.

Kovács, A., & Mehler, J. (2009a). Cognitive gains in 7-month-old bilingual infants. Proceedings of the National Academy of Sciences, 106, 6556-6560.

Kovács, A., & Mehler, J. (2009b). Flexible learning of multiple speech structures in bilingual infants. Science, 325, 611-612.

Kovelman, I., Shalinsky, M. H., White, K. S., Schmitt, S. N., Berens, M. S., Paymer, N., & Petitto, L. A. (2009). Dual Language use in sign-speech bimodal bilinguals: fNIRS brain-imaging evidence. Brain and Language, 109, 112-123.

Kuczaj, S. A. (1978). Children's judgements of grammatical and ungrammatical irregular past-tense verbs, Child Development, 49, 319-326.

Levelt, C. C. (1994). On the acquisition of place. Ph.D. dissertation. Leiden University, HIL Dissertation Series 8.

Lew-Williams, C., & Fernald, A. (2007). Young children learning Spanish made rapid use of the grammatical gender in spoken word recognition. Psychological Science, 18, 193-198.

MacWhinney, B. (2000). The CHILDES project: Tools for analyzing talk (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Mani, N. (to appear.) Phonological Acquisition. In B. Botma, N. Kula, & K. Nasukawa. (Eds.), Continuum Handbook to Phonology. Amsterdam: Continuum.

Marcus, G., Pinker, S., Ullman, M., Hollander, M., Rosen, J., & Xu, F. (1992). Overregularization in Language Acquisition. Monographs of the Society for Research in Child Development, 57(4).

Mattys, S. L. & Jusczyk, P. W. (2001). Phonotactic cues for segmentation of fluent speech by infants. Cognition, 78, 91-121

McMurray, B., & Aslin, R. N. (2004). Anticipatory eye movements reveal infants' auditory and visual categories. Infancy, 6, 203-229.

McMurray, B., Spivey, M., and Aslin, R. N. (2000). The Perception of Consonants by Adults and Infants: Categorical or Categorized? Preliminary Results. In K. Crosswhite, & J. Magnuson (Eds.,) The University of Rochester Working Papers in the Language Sciences (pp. 215-256).

Meek, J. (2002). Basic principles of optical imaging and application to the study of infant development. Developmental Science, 5, 371-380.

Mehler J., Nespor, M., Gervain J., Endress A. D., & Shukla M., (2008). Mechanisms of language acquisition: imaging and behavioral evidence. In: C. A. Nelson & M. Luciana (eds.): Handbook of Developmental Cognitive Neuroscience. 2nd edition (pp. 325-336). Cambridge, Mass.: MIT Press.

Munson, B., Kurtz, B. A., & Windsor, J. (2005). The influence of vocabulary size, phonotactic probability, and wordlikeness on nonword repetitions of children with and without specific language impairment. Journal of Speech, Language, and Hearing Research, 48, 1033–1047.

Nation, K., & Marshall, C. M., & Altmann, G. T. (2003). Investigating individual differences in children's real-time sentence comprehension using language-mediated eye movements. Journal of Experimental Child Psychology, 86, 314-329.

Ohala, D. (1999). The influence of sonority on children's clusters reductions. Journal of Communication Disorders, 32, 397-422.

Ohala, D. (2008). Phonological acquisition in a first language. In M. Zampini and J. Hansen (Eds.), Phonology and second language acquisition (pp. 19-29). John Benjamins.

Parisse, C., & Soubeyrand, E. (2003). Computerized test for oral language perception: TIPLO. Journal Of Multilingual Communication Disorders, 1, 53-61.

Peña, M., Maki, A., Kovačić, D., Dehaene-Lambertz, G., Koizumi, H., Bouquet, F., & Mehler, J. (2003). Sounds and silence: an optical topography study of language recognition at birth. Proceedings of the National Academy of Sciences, 100, 11702-5.

Rose, Y., MacWhinney, B., Byrne, R., Hedlund, G., Maddocks, K., O'Brien, P., & Wareham, T. (2006). Introducing Phon: A Software Solution for the Study of Phonological Acquisition. In D. Bamman, T. Magnitskaia, & C. Zaller (Eds.) Proceedings of the 30th Annual Boston University Conference on Language Development (pp. 489-500). Somerville, MA: Cascadilla Press.

Scobbie, J. M., Gibbon, F., Hardcastle, W. F., & Fletcher, P. (2000). Covert contrast as a stage in the acquisition of phonetics and phonology. In M. Broe and J. Pierrehumbert (Eds.) Papers in Laboratory Phonology V: Language Acquisition and the Lexicon, 194-207. Cambridge: Cambridge University Press.

Shvachkin, N. K. (1973). The development of phonemic speech perception in early childhood. In C. Ferguson & D. Slobin (Eds.), Studies of child language development (pp. 91-127). New York : Holt, Rinehart & Winston. (Original work published 1948).

Singh, L., Morgan, J., & White, K. (2004). Preference and processing: the role of speech affect in early spoken word recognition. Journal of Memory and Language, 51, 173-189.

Smith, N. V. (1973). The acquisition of phonology: A case study. Cambridge: Cambridge University Press.

Song, J. Y., Sundara, M., & Demuth, K. (2009). Effects of phonology on children's production of English 3rd person singular –s. Journal of Speech, Language, and Hearing Research, 52, 623-642.

Stoel-Gammon, C. (1985). Phonetic inventories, 15-24 months: a longitudinal study. Journal of Speech and Hearing Research, 28, 505–512

Stoel-Gammon, C., & Sosa, A. V. (2006). Phonological Development. In E. Hoff & M. Schatz, Blackwell Handbook of Language Development, pp. 238-256. John Wiley & Sons.

Storkel, H. L. (2004). Methods for minimizing the confounding effects of word length in the analysis of phonotactic probability and neighborhood density. Journal of Speech, Language, and Hearing Research, 47, 1454-1468.

Sundara, M., Demuth, K. & Kuhl, P. K. (2008, Nov.) Positional effects on 2-year-olds' comprehension and production of 3rd person singular –s. Poster presented at the 33rd Boston University Conference on Language Development, Boston, MA.

Swingley, D., & Aslin, R. N. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. Psychological Science, 13, 480-484.

Treiman, R., & Danis. C. (1988). Syllabification of intervocalic consonants. Journal of Memory and Language, 27, 87-104.

Tyler, M. D., Tyler, L., & Burnham, D. K. (2005). The delayed trigger voice key: An improved analogue voice key for psycholinguistic research. Behavior Research Methods, 37, 139-147.

Van Heugten, M., & Johnson E. K. (2009). The robustness of infants' early word representations. Journal of the Canadian Acoustical Association, 37, 148-149.

Vihman, M. M., DePaolis, R. A., & Keren-Portnoy, T, (2009). Babbling and words: A Dynamic Systems perspective on phonological development. In E. L. Bavin (Ed.) The Cambridge Handbook of Child Language (pp.163-182). Cambridge: Cambridge University Press.

Walley, A. C. (2005). Speech perception in childhood. In D. B. Pisoni & R. E Remez (Eds.), Handbook of speech perception (pp.449-468). Oxford: Blackwel

Werker, J. F., & Curtin, S. (2005). PRIMIR: A Developmental Framework of Infant Speech Processing. Language Learning and Development, 1, 197-234.

Zamuner, T. S. (2003). Input-Based Phonological Acquisition. New York, NY: Routledge.

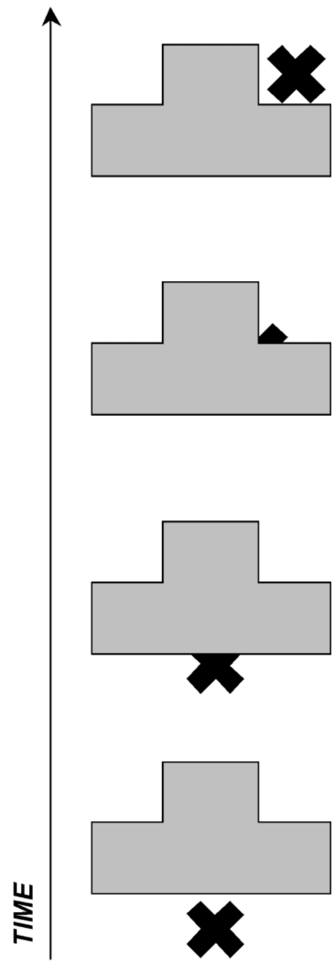
Zamuner, T. S. (2009). Phonotactic probabilities at the onset of language development: Speech production and word position. Journal of Speech, Language, and Hearing Research, 52, 49-60.

Zamuner, T. S., Gerken, L.A., & Hammond, M. (2004). Phonotactic probabilities in young children's speech production. Journal of Child Language, 31, 515-536.

Zamuner, T.S., & Ohala, D. (1999). Preliterate children's syllabification of intervocalic consonants. In A. Greenhill, H. Litterfield, & C. Tano (Eds.), Proceedings of the 23rd Boston University Conference on Language Development Vol. 2. (pp. 753-763). Somerville, MA: Cascadilla Press.

Zamuner, T. S., Kerkhoff, A., & Fikkert, P. (2008). Phonotactics and morpho-phonology in early child language: Evidence from Dutch. Manuscript under review.

Figure 1: The time course of a single trial: Visual stimuli start below the occluder, move behind it, and then emerge on either the right or left. Taken from McMurray and Aslin (2004).



---

[1] Studies show that sound-changes in children's phonological system depend on many factors, such as the type of sound change, word position and word frequency (Gierut, 2001; Gierut and Storkel, 2002).

[2] Note that by adult-like, we do not mean 'perfect perception'. Adults perception is not perfect or free of errors. Adults are known to make the occasional 'slip of the ear' (e.g. Bond, 1999) as well as the occasional 'slip of the tongue' (Cutler, 1982).

[3] Two primary methods used to evaluate children's productions are the independent analysis and relational analysis. Independent analyses measure children's productions independent of the adult target without consideration of whether children produce the adult target correctly. Relational analyses measure children's productions as they relate to the adult target form (Stoel-Gammon, 1985; Stoel-Gammon and Sosa, 2006).